Astronomical Data Analysis Software and Systems VI
ASP Conference Series, Vol. 125, 1997
Gareth Hunt and H. E. Payne, eds.

Time Series Analysis of Unequally Spaced Data: Intercomparison Between Estimators of the Power Spectrum

V. V. Vityazev

Astronomy Department, St. Petersburg University, Bibliotechnaya pl.2, Petrodvorets, St. Petersburg, 198904, Russia.

Abstract. It is shown that the likeness of the periodogram and the LS-spectrum (both estimators of the power spectrum are widely used in the spectral analysis of time series), depends on the properties of the spectral window $W(\omega)$ corresponding to the distribution of time points. The main results are: a) all the estimators evaluated at frequency ω are identical if $W(2\omega) = 0$; b) the Schuster periodogram differs from the LS-spectra at the frequencies $\omega = \hat{\omega}_k/2$, where $\hat{\omega}_k$ are the frequencies at which the spectral window has large side peaks due to irregular distribution of time points. Two examples for situations typical in astronomy illustrate these conclusions.

1. Introduction

In various branches of astronomy, we face the problem of finding periodicities hidden in observations. If the data are regularly spaced in time, the Schuster periodogram is the basic tool for evaluating the power spectra (Marple 1987; Terebizh 1992). Unfortunately, astronomical observations are irregular for various reasons: day-time changes, weather conditions, positions of the object under observations, etc. Present day theory and practice of the spectral analysis of unequally spaced time series are based on two approaches. The first one employs the Schuster periodogram (Deeming 1975; Roberts et al. 1987). The second one uses the procedure of the least squares fitting of a sinusoid to the data (Barning 1962; Lomb 1976; Scargle 1982) with resulting estimators known as the LS-spectra. The most valuable feature of the LS-spectra is well defined statistical behavior. At the same time, the LS-spectra lose very important properties: description in terms of the spectral window, connection with the correlation function, etc. On the other hand, the Schuster periodogram of a gapped time series satisfies all the fundamental relations of the classical spectral analysis, but its statistical properties are complicated as compared to the case of regular data. It is worth mentioning that despite different theoretical foundations, the Schuster periodogram and the LS-spectra frequently turn out to be almost identical. This similarity requires an explanation, and we are trying to find situations when the Schuster periodogram and the LS-spectra are very close to each other or differ greatly. The ultimate goal of this study is to clarify the properties of various techniques which are used to derive the periodicities from the unequally distributed data.

2. Two Estimators of the Power Spectrum

For a set of N observations $x_k = x(t_k)$, $k = 0, 1, \ldots, N-1$ with zero mean obtained at arbitrary times t_k, we can set up the model

$$f(t) = \sum_{i=1}^{2} a_i \phi_i(t), \tag{1}$$

where

$$\phi_1(t) = \cos \omega t, \qquad \phi_2(t) = \sin \omega t. \tag{2}$$

Using the following notation

$$(p, q) = \frac{1}{N} \sum_{k=0}^{N-1} p(t_k) q(t_k), \qquad \|p\|^2 = (p, p), \tag{3}$$

the "classical" estimator of the power spectrum (the Schuster periodogram) can be written in the form

$$S(\omega) = (x, \phi_1)^2 + (x, \phi_2)^2 = \frac{1}{N^2} \ \mid \sum_{k=0}^{N-1} x_k e^{-i\omega t_k} \mid^2 . \tag{4}$$

If the signal contains a sine function of frequency ω_0, then the product $x_k e^{-i\omega t_k}$ makes a large contribution to S provided that $\omega = \omega_0$. In other words, the Schuster periodogram, to the limit of normalizing factor, is a square of the correlation coefficient between the data and a harmonic function.

The alternative estimator of the power spectrum based on the least squares fitting of the sine function to the data was proposed by Lomb (1976) and Scargle (1982). Their approach is based on the introduction of the new time points

$$\hat{t_k} = t_k - \frac{1}{2\omega} \ \arctan \frac{\sum_k \sin 2\omega t_k}{\sum_k \cos 2\omega t_k}, \tag{5}$$

where the time shift provides the orthogonality of the functions

$$\hat{\phi}_1(t) = \cos \omega \hat{t_k}, \qquad \hat{\phi}_2(t) = \sin \omega \hat{t_k}. \tag{6}$$

Under this assumption the LS-spectrum looks as follows:

$$L(\omega) = \frac{1}{2} \left[\frac{(x, \hat{\phi}_1)^2}{\|\hat{\phi}_1\|^2} + \frac{(x, \hat{\phi}_2)^2}{\|\hat{\phi}_2\|^2} \right]. \tag{7}$$

Thus we see that the Schuster periodogram differs from the LS-spectra by definition.

The intercomparison between the Schuster periodogram and the LS-spectrum is given by the following

Theorem. *At the set of frequencies that satisfy equation*

$$W(2\omega) = 0, \tag{8}$$

where the spectral window $W(\omega)$ is

$$W(\omega) = \frac{1}{N^2} \mid \sum_{k=0}^{N-1} e^{-i\omega t_k} \mid^2 \tag{9}$$

the Schuster periodogram and the LS-spectra are identical.

3. The Spectral Windows for Typical Distribution of Time Points

In this section we consider two typical distributions of points for which the frequencies that satisfy Eq. (8) do exist.

3.1. Time Series with Periodic Gaps

Astronomical observations are often performed with periodic gaps. Ground-based observations are interrupted by day-night alternation; the observations from a space vehicle are usually stopped when the satellite enters the radiation belts. To model the situations we suppose that, in the set of observations with a constant interval Δt, one has n successive observations and p successive missing points, and the group of $n+p$ points is repeated m times, so the period of gaps is $\Delta T = (n+p)\Delta t$. In the previous paper (Vityazev 1994) it was shown that in this case the spectral window looks as follows

$$W(\omega) = \frac{\sin^2(n\omega\Delta t/2)}{n^2 \sin^2(\omega\Delta t/2)} \, \frac{\sin^2(m\omega\Delta T/2)}{m^2 \sin^2(\omega\Delta T/2)}. \tag{10}$$

It is easy to find that the frequencies

$$\omega_j = \frac{2\pi}{m\Delta T} j, \quad j = 1, 2, ... \tag{11}$$

satisfy Eq. (8), provided that $j \neq m/2, m, ...$ if m is even and $j \neq m, 2m, ...$, if m is odd.

3.2. Observations with a Long Gap

Considered here is a situation where two sets of observations (each one consisting of n successive points) are separated by p missing points forming the gap. As earlier, all the points are supposed to be regularly spaced over the time interval $\Delta t = const$. Now, for the spectral window we have (Vityazev 1994)

$$W(\omega) = \frac{\sin^2(n\omega\Delta t/2)}{n^2 \sin^2(\omega\Delta t/2)} \, [\cos((n+p)\omega\Delta t)]/2. \tag{12}$$

It is not difficult to show that the frequencies

$$\omega_j = \frac{\pi}{(n+p)\Delta t} \, (j + \frac{1}{2}), \; j = 0, 1, \; ... \; , n+p-1, \tag{13}$$

satisfy the condition of Eq. (8).

4. Conclusions

The LS-spectra gained popularity due to the fact that they retain the exponential distribution of their accounts when the time series is assumed to be white noise. Now we see that at frequencies that satisfy Eq. (8) the Schuster periodogram retains the exponential distribution too.

The Schuster periodogram differs from the LS-spectra only at the frequencies that satisfy the condition $1 - W(2\omega) \ll 1$. It means that the discrepancies between the Schuster periodogram and the LS-spectra are large when the time series contain a harmonic of the frequency, the double value of which coincides with the frequency at which the spectral window has a large side peak. In the case of periodical gaps it happens when the period of a signal hidden in the data is one half the period of the gaps. In this situation (Vityazev 1997a), the spectral estimation faces unrealistic intensities of the spectral peaks and the strong dependence of the heights of peaks on the phase of the signal. It is very important to emphasize that these problems come not from the choice of the tool to evaluate the power spectrum; they originate from mixing two sources of the periodicities: one is the physical process that we observe and another one is a periodical interruption of observations. In astronomy, the rotation and revolution of the Earth impose diurnal and annual cycles on the Earth-based observations. The periods hidden in observations of the Sun, stars, quasars, etc., are hardly connected physically with the periods specific to the Earth. For these observations, the probability of mixing periodicities is negligible. On the contrary, if we study the Earth from the Earth (such is the case with astrometric observations of the Earth's rotation parameters), then the semi-annual period in the Earth's rotation interferes with the annual gaps in observations.

For further details the reader is referred to Vityazev (1997a, 1997b).

References

Barning, F. J. M. 1962, B.A.N., 17, N1, 22

Deeming, T. J. 1975, Ap&SS, 36, 137

Lomb, N. R. 1976, Ap&SS, 39, 447

Marple, Jr., S. L. 1987, Digital Spectral Analysis with Applications (Englewood Cliffs, NJ: Prentice-Hall)

Roberts, D. H., Lehar, J., & Dreher, J. W. 1987, AJ, 93, 968

Scargle, J. D. 1982, ApJ, 263, 835

Terebish, V. Yu. 1992. Time Series Analysis in Astrophysics (Moscow: Nauka)

Vityazev, V. V. 1994, Astron. and Astrophys. Tr., 5, 177

Vityazev, V. V. 1997a, A&A, in press

Vityazev, V. V. 1997b, A&A, in press